

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE  
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of	Atty. Docket
FABIO VIGNOLI	NL 021053
	Confirmation No. 1612
Serial No. 10/532,469	Group Art Unit: 2626
Filed: APRIL 22, 2005	Examiner: SAINT CYR, L.
Title: CONTROLLING AN APPARATUS BASED ON SPEECH	

Mail Stop Appeal Brief-Patents  
Board of Patent Appeals and Interferences  
United States Patent and Trademark Office  
P.O. Box 1450  
Alexandria, VA 22313-1450

**APPEAL BRIEF**

Sir:

Appellants herewith respectfully present a Brief on Appeal as follows, having filed a Notice of Appeal on April 21, 2009:

REAL PARTY IN INTEREST

The real party in interest in this appeal is the assignee of record Koninklijke Philips Electronics N.V., a corporation of The Netherlands having an office and a place of business at Groenewoudseweg 1, Eindhoven, Netherlands 5621 BA.

RELATED APPEALS AND INTERFERENCES

Appellants and the undersigned attorney are not aware of any other appeals or interferences which will directly affect or be directly affected by or having a bearing on the Board's decision in the pending appeal.

STATUS OF CLAIMS

Claims 1-19 are pending in this application. Claims 1-19 are rejected in the Final Office Action mailed in January 21, 2009. This rejection was upheld, in an Advisory Action that mailed April 9, 2009. Claims 1-19 are the subject of this appeal.

STATUS OF AMENDMENTS

Appellants filed on March 18, 2009 an after final amendment in response to a Final Office Action mailed on January 21, 2009. The after final amendment did not include any amendments. In an Advisory Action mailed on April 9, 2009, it is indicated that the after final amendment filed on March 18, 2009 does not place the application in condition for allowance. This Appeal Brief is in response to the Final Office Action mailed January 21, 2009, that finally rejected claims 1-19, which remain finally rejected in the Advisory Action mailed on April 9, 2009.

SUMMARY OF THE CLAIMED SUBJECT MATTER

The present invention, for example, as recited in independent claim 1 and shown in FIGs 1-2, as well as described on page 6, lines 1-6 of the specification, is directed to a speech control unit 100 for controlling an apparatus 200 on basis of speech, comprising a microphone array, comprising multiple microphones 102, 104, 106, 108, 110 for receiving respective audio signals 103, 105, 107, 109, 111.

As shown in FIG 1, and described on page 6, lines 7-14, speech control unit 100 includes a beam forming module 116 for extracting a speech signal 117 of a user U1, from the audio signals 103, 105, 107, 109, 111 as received by the microphones 102, 104, 106, 108, 110, by enhancing first components of the audio signals which represent an utterance originating from a first position P1 of the user U1 relative to the microphone array. A speech recognition unit 118 creates an instruction for the apparatus 200 based on recognized speech items of the speech signal. Further, a keyword recognition system 120 recognizes a predetermined keyword that is

spoken by the user U1 and which is represented by a particular audio signal.

As described on page 3, lines 22-29, and page 6, line 10 to page 7, line 3 of the specification, the speech control unit 100 is arranged to control the beam forming module 116, on basis of the recognition of the predetermined keyword, in order to enhance second components of the audio signals which represent a subsequent utterance originating from a second position P2 of the user U1 relative to the microphone array. The recognition of the predetermined keyword at the second position calibrates the beam forming module 116 to follow the user from the first position P1 to the second position P2 so that the subsequent utterance originating from the second position P2 are accepted while utterances of other users U2 at other positions are discarded. The second position P2 includes orientation and distance relative to the microphone array, as described on page 4, lines 22-28. Further, the speech control unit 100 is also configured to discriminate between sounds originating from users who are located in front of each other. The subsequent utterance originating from the second position P2 will

be discarded if not preceded by the recognition of the predetermined keyword originating from the second position P2, as described on page 3, line 29 to page 4, line 3.

The present invention, for example, as recited in independent claim 9 and shown in FIGs 1-2, as well as described on page 6, lines 1-6 of the specification, is directed to a method of controlling an apparatus 200 on basis of speech, comprising receiving respective audio signals 103, 105, 107, 109, 111 by a microphone array, comprising multiple microphones 102, 104, 106, 108, 110.

As shown in FIG 1, and described on page 6, lines 7-14, the method further includes extracting a speech signal 117 of a user U1, from the audio signals 103, 105, 107, 109, 111 as received by the microphones 102, 104, 106, 108, 110, by enhancing first components of the audio signals which represent an utterance originating from a first position P1 of the user relative to the microphone array; and recognizing a predetermined keyword that is spoken by the user U1 based on a particular audio signal.

As described on page 3, lines 22-29, and page 6, line 10 to



page 7, line 3 of the specification, the extraction of the speech signal of the user U1 is controlled on basis of the recognition of the predetermined keyword, in order to enhance second components of the audio signals which represent a subsequent utterance originating from a second position P2 of the user U1 relative to the microphone array, while discarding utterances of other users at other positions. The second position P2 includes an orientation and a distance relative to the microphone array so that sounds originating from users who are located in front of each other are discriminated, as described on page 4, lines 22-28 of the specification.

As described on page 3, line 29 to page 4, line 3, an instruction for the apparatus is created based on recognized speech items of the speech signal, and the subsequent utterance originating from the second position P2 is discarded if not preceded by the recognition of the predetermined keyword originating from the second position P2.

GROUND OF REJECTION TO BE REVIEWED ON APPEAL

Whether claims 1-19 of U.S. Patent Application Serial No.  
10/532,469 are unpatentable under 35 U.S.C. §103(a) over U.S.  
Patent No. 7,136,817 (Schroder) in view of U.S. Patent No.  
7,050,971 (Kaufholz) and U.S. Patent Application Publication No.  
2002/0181723 (Kataoka).

ARGUMENT

Claims 1-19 are is said to be unpatentable over Schroder,  
Kaufholz and Kataoka.

Appellants respectfully request the Board to address the patentability of independent claims 1 and 9, and further claims 2-8 and 10-19 as depending from independent claims 1 and 9, based on the requirements of independent claims 1 and 9. This position is provided for the specific and stated purpose of simplifying the current issues on appeal. However, Appellants herein specifically reserve the right to argue and address the patentability of claims 2-8 and 10-19 at a later date should the separately patentable subject matter of claims 2-8 and 10-19 later become an issue. Accordingly, this limitation of the subject matter presented for appeal herein, specifically limited to discussions of the patentability of independent claims 1 and 9 is not intended as a waiver of Appellants' right to argue the patentability of the further claims and claim elements at that later time.

As correctly noted on page 4 of the Final Office Action,

Schroder and Kaufholz do not disclose or suggest that the second position includes "an orientation and a distance relative to the microphone array, and the speech control unit being configured to discriminate between sounds originating from users who are located in front of each other," as recited in independent claim 1, and similarly recited in independent claim 9. Paragraphs [0006] and [0047] of Kataoka are cited in an attempt to remedy the deficiencies in Schroder and Kaufholz.

Kataoka is directed to a robot where the attention direction of a robot can be aligned with a directivity direction of a microphone array. Paragraph [0006] of Kataoka merely describes that the robot recognize voices from an arbitrary direction, where the direction of a sound source is estimated by a microphone array. Further, paragraph [0047] merely describes that the attention direction of the robot is aligned with an incident direction of the voice, and that voices from the attention direction (the directivity direction of the microphone array) are emphasized.

It is respectfully submitted that Schroder, Kaufholz, Kataoka, and combination thereof, do not teach or suggest the present

invention as recited in independent claim 1, and similarly recited in independent claim 9 which, amongst other patentable elements, recites (illustrative emphasis provided):

wherein the recognition of the predetermined keyword at the second position calibrates the beam forming module to follow the user from the first position to the second position so that the subsequent utterance originating from the second position are accepted while utterances of other users at other positions are discarded, the second position including an orientation and a distance relative to the microphone array, and the speech control unit being configured to discriminate between sounds originating from users who are located in front of each other.

Discriminating between sounds originating from users who are located in front of each other is nowhere disclosed or suggested in Schroder, Kaufholz and Kataoka, alone or in combination. Kataoka is merely concerned with direction, and NOT distance, and does not disclose or suggest discriminate between sounds originating from users who are located in front of each other, as recited in independent claims 1 and 9. Further, paragraphs [0007] and [0047] of Kataoka, noted in the Advisory Action, merely "directional estimation of a sound source" as specifically recited in paragraph [0007], line 3. Further, as noted above, paragraph [0047] of

Kataoka describes emphasizing voices from the attention direction (i.e., the directivity direction of the microphone array), "whereby the recognition rate of the voices from that direction can be enhanced." (Paragraph [0047], last sentence) Such recitations of paragraphs [0007] and [0047] have nothing to do with distance and are merely concerned with direction.

Accordingly, it is respectfully submitted that independent claims 1 and 9 are allowable. In addition, it is respectfully submitted that claims 2-8 and 10-19 are also allowable at least based on their dependence from independent claims 1 and 9.


In addition, Appellants deny any statement, position or averment of the Examiner that is not specifically addressed by the foregoing argument and response. Any rejections and/or points of argument not addressed would appear to be moot in view of the presented remarks. However, the Appellants reserve the right to submit further arguments in support of the above stated position, should that become necessary. No arguments are waived and none of the Examiner's statements are conceded.

CONCLUSION

Claims 1-19 are patentable over Schroder, Kaufholz and  
Kataoka.

Thus, the Examiner's rejections of claims 1-19 should be  
reversed.

Respectfully submitted,

By   
Dicran Halajian, Reg. 39,703  
Attorney for Appellants  
June 22, 2009

**THORNE & HALAJIAN, LLP**  
Applied Technology Center  
111 West Main Street  
Bay Shore, NY 11706  
Tel: (631) 665-5139  
Fax: (631) 665-5101

## CLAIMS APPENDIX

1. ( Previously Presented) A speech control unit for controlling an apparatus on basis of speech, comprising:

a microphone array, comprising multiple microphones for receiving respective audio signals;

a beam forming module for extracting a speech signal of a user, from the audio signals as received by the microphones, by means of enhancing first components of the audio signals which represent an utterance originating from a first position of the user relative to the microphone array;

a speech recognition unit for creating an instruction for the apparatus based on recognized speech items of the speech signal;  
and

a keyword recognition system for recognition of a predetermined keyword that is spoken by the user and which is represented by a particular audio signal;

the speech control unit being arranged to control the beam forming module, on basis of the recognition of the predetermined



keyword, in order to enhance second components of the audio signals which represent a subsequent utterance originating from a second position of the user relative to the microphone array;

wherein the recognition of the predetermined keyword at the second position calibrates the beam forming module to follow the user from the first position to the second position so that the subsequent utterance originating from the second position are accepted while utterances of other users at other positions are discarded, the second position including an orientation and a distance relative to the microphone array, and the speech control unit being configured to discriminate between sounds originating from users who are located in front of each other; and

wherein the subsequent utterance originating from the second position will be discarded if not preceded by the recognition of the predetermined keyword originating from the second position.

2.(Previously Presented) The speech control unit as claimed in claim 1, wherein the keyword recognition system is arranged to recognize the predetermined keyword that is spoken by another user

and the speech control unit being arranged to control the beam forming module, on basis of this recognition, in order to enhance third components of the audio signals which represent another utterance originating from a third position of the other user relative to the microphone array.

3.(Previously Presented) The speech control unit as claimed in claim 1, wherein a first one of the microphones of the microphone array is arranged to provide the particular audio signal to the keyword recognition system.

4.(Previously Presented) The speech control unit as claimed in claim 1, wherein the beam forming module is arranged to determine a first position of the user relative to the microphone array.

5.(Original) An apparatus comprising:  
a speech control unit for controlling the apparatus on basis of speech as claimed in claim 1; and

processing means for execution of the instruction being created by the speech control unit.

6. (Previously Presented) The apparatus as claimed in claim 5, the apparatus arranged to show that the predetermined keyword has been recognized.

7. (Previously Presented) The apparatus as claimed in claim 6, further comprising audio generating means for generating an audio signal in order to show that the predetermined keyword has been recognized.

8. (Original) A consumer electronics system comprising the apparatus as claimed in claim 5.

9. (Previously Presented) A method of controlling an apparatus on basis of speech, comprising the acts of:

receiving respective audio signals by means of a microphone array, comprising multiple microphones;

extracting a speech signal of a user, from the audio signals as received by the microphones, by means of enhancing first components of the audio signals which represent an utterance originating from a first position of the user relative to the microphone array;

recognizing a predetermined keyword that is spoken by the user based on a particular audio signal and controlling the extraction of the speech signal of the user, on basis of the recognition of the predetermined keyword, in order to enhance second components of the audio signals which represent a subsequent utterance originating from a second position of the user relative to the microphone array while discarding utterances of other users at other positions, the second position including an orientation and a distance relative to the microphone array so that sounds originating from users who are located in front of each other are discriminated;

creating an instruction for the apparatus based on recognized speech items of the speech signal; and

discarding the subsequent utterance originating from the

second position if not preceded by the recognition of the predetermined keyword originating from the second position.

10.(Previously Presented) The speech control unit of claim 1, wherein the user is informed by indications that the speech control unit is not active, is in an active state and ready to receive the utterance, or is in a state of calibration.

11.(Previously Presented) The speech control unit of claim 10, wherein the indications include an animal in a sleeping state indicating that the speech control unit is not active, and in an awake state indicating that the speech control unit is in the active state.

12.(Previously Presented) The speech control unit of claim 11, wherein progress of the active state is indicated by an angle of ears of the animal.

13.(Previously Presented) The speech control unit of claim

12, wherein the ears are fully raised at a beginning of the active state, and fully down at an end of the active state.

14. (Previously Presented) The speech control unit of claim 11, wherein the animal has an understanding look when the utterance is recognized and a puzzled look when the utterance is not recognized.

15. (Previously Presented) The method of claim 9, further comprising the act of informing the user by indications that the apparatus is not active, is in an active state and ready to receive the utterance, or is in a state of calibration.

16. (Previously Presented) The method of claim 15, wherein the indications include an animal in a sleeping state indicating that the speech control unit is not active, and in an awake state indicating that the speech control unit is in the active state.

17. (Previously Presented) The method of claim 16, wherein

progress of the active state is indicated by an angle of ears of the animal.

18. (Previously Presented) The method of claim 17, wherein the ears are fully raised at a beginning of the active state, and fully down at an end of the active state.

19. (Previously Presented) The method of claim 16, wherein the animal has an understanding look when the utterance is recognized and a puzzled look when the utterance is not recognized.

**EVIDENCE APPENDIX**

None



**RELATED PROCEEDINGS APPENDIX**

None